African Journal of Mathematics and Statistics Studies

ISSN: 2689-5323

Volume 8, Issue 3, 2025 (pp. 47-65)

www.abjournals.org

SEVEN-FACTOR CENTRAL COMPOSITE DESIGN ROBUST TO A PAIR OF MISSING OBSERVATIONS

Okon Emmanuel A.*, Ugbe Thomas A., and Akpan Stephen S.

Department of Statistics, University of Calabar, Calabar.

*Corresponding Author's Email: ugbethomas1@gmail.com

Cite this article:

Okon Emmanuel A., Ugbe Thomas A., Akpan Stephen S. (2025), Seven-Factor Central Composite Design Robust to A Pair of Missing Observations. African Journal of Mathematics and Statistics Studies 8(3), 47-65. DOI: 10.52589/AJMSS-OO1VGAGD

Manuscript History

Received: 5 Feb 2025

Accepted: 7 Mar 2025

Published: 25 Jul 2025

ABSTRACT: *Seven-factor central composite design has been studied with respect to a pair of missing observations. The central composite designs comprise of factorial, axial and center parts. All possible combinations of the missing points are considered. The study is based on the criterion of minimizing the maximum loss, which depends mainly on the number of factors involved in the experiment, the distance of the axial point from the design centre (〈) and position of the missing points. The work considered the case of minimizing the maximum loss when two observations are missing in a Central Composite Design (CCD) with k =7. The loss of every possible combination of two missing observations was calculated using minimax loss criterion, and groups of two missing observations producing the same losses formed. The process was repeated for a range of 〈 values to locate the 〈 for which the maximum loss is minimum. It was discovered that the loss effect of missing a pair of factorial points is a decreasing function of increasing 〈, while the loss effect of a pair of axial points is a decreasing and increasing function of increasing 〈. The loss effect of missing a factorial and axial point has no specific direction of increase or decrease on increasing 〈 values. It was also discovered that irrespective of the value of k, when a couple of a pair of observations is missing the design will break down.*

KEYWORDS: Central composite design, Minimax loss criterion, Axial points, Orthogonality, Factorial points, outlier.

INTRODUCTION

In a practical experiment, situations may arise where some observations are lost or destroyed or unavailable due to cost, time or for certain other reasons that are beyond the control of the experimenter. Unavailability of the observations destroys the orthogonality and the balance of the design and also affects the analysis of variance. Missing observations being an unavoidable problem also complicate the statistical analysis of data, thereby affecting the results of a designed experiment quite badly. The problem becomes more serious especially in case of high cost experiments, where repeating the experiment becomes almost impossible.

Missing observations or outliers can occur in any experiment irrespective of how well the experimenter had planned the experimental runs. Sequel to the existence of missing observations in an experiment; the estimation of all the parameters in the response surface model may not be possible. The inestimable model parameters are worst in a poorly planned experiment, because the loss of a small number of observations or outliers may yield a singular information matrix, where no parameters could be estimated. Consequently, it is necessary to set-up designs that are robust to the unavailability of a set of observations, that is, if for any given reasons some data were not collected in the design; it would still be possible to estimate all the parameters in the model. Akhtar and Prescott (1986) developed minimax loss criterion based on D-optimality and used it to examine effects of one or two missing values in a second order model response surface design for $2 \leq k \leq 6$. They further compared the selected designs which are robust to one or two missing observations with rotatable designs and the outlier robust designs.

Akhtar (2001) had studied the structure of missing one and two observations for the following configurations of ccd for $k = 5$;

(i) design with half replicate of factorial part $(\frac{1}{2}F + A)$

(ii) design with one replication of factorial and axial part $(F + A)$

(iii) design with one replication of factorial and two replication of axial parts $(F + 2A)$.

Akram (2002) investigated ccd robust to three missing observations for $2 \leq k \leq 6$ and compared the results with designs of the same configurations, such as cuboidal, spherical, and orthogonal, among others. Akram *et al*. (2003) studied the structure of loss-sensitive combinations of missing observations in central composite designs. They carried out computations in terms of eigenvalues for the comparison of the class of $ccds$ on the basis of E-optimality. They found that due to missing a combination of three observations, the increase in variance of the estimates is less as the number of design points increases regardless of whether the missing combination consists of factorial, axial or centre point. It was also discovered in their investigation that the missing of an axial point may create more problem than the missing of a factorial point when measuring the variance for $k \geq 4$.

Other recent studies on missing observations are as follows. Ossai and Oladugba (2018) derived an algebraic expressions for estimating missing data when one or more observation(s) are missing in rectangular lattice designs with repetition using the method of minimizing the residual sum of squares. Results showed that the estimated value(s) were significantly approximate to that of the actual value(s); Fujiwara and Matsuura (2020) derived a

Article DOI: 10.52589/AJMSS-OO1VGAGD
DOI URL: https://doi.org/10.52589/AJMSS-OO1VGAGD

mathematical expression for the inflation amount of the prediction variance. It turns out that, for rotatable central composite designs, the inflation amount of the prediction variance depends only on the euclidean norms and the inner product of the two vectors of factor values at which the observation is missing and the response is predicted; Rashid, Akbar and Arshad (2022) investigated the impact of a missing observation on the estimation and prediction capabilities as well as on the relative A-, D- and G-efficiencies of augmented Box-Behnken designs. They observed that precision of model parameter estimates as well as relative A-, D- and G-efficiencies are less affected by missing a center run, but more by missing other design points; Yankam and Oladugba (2023) constructed new orthogonal uniform minimax loss designs. Their designs compared with central composite designs, small composite designs, orthogonal array composite designs, orthogonal array composite minimax loss designs and orthogonal uniform composite designs based on the D-efficiency, A-efficiency and T-efficiency for estimating the parameters of the second-order model and under the generalized scaled standard deviations.

This work is focused on setting-up a design that is robust to a pair of missing observations. It is also extended to the comparison of the obtained robust design with designs of the same configuration in the literature, such as rotatable central composite designs, face-centered central composite designs, orthogonal central composite designs, outlier central composite designs and spherical central composite designs.

METHODOLOGY

Second-Order Central Composite Design Model

We shall focus our attention on second-order central composite design using the normal coded values of (± 1). The general form of the second-order model with k factors is given as:

$$\hat{y} = \hat{\beta}_0 + \sum_{i=1}^{k} \hat{\beta}_i x_i + \sum_{i=1}^{k} \hat{\beta}_{ii} x_i^2 + \sum_{i<j} \hat{\beta}_{ij} x_i x_j \tag{1}$$

This equation may be written in matrix notation

$$\hat{y} = X\underline{\hat{\beta}} + \underline{\varepsilon} \tag{2}$$

where $\underline{y}$ is a $n \times 1$ column vector of predicted responses corresponding to n trials. $\underline{\hat{\beta}}$ is a $p \times 1$ column vector of estimates of parameters. X is a $n \times p$ design model matrix.

For second-order response surface model, the number of coefficient is given by $p = \binom{k+2}{2} = \frac{(k+1)(k+2)}{2}$

That is, the quadratic model has a constant term, k, linear terms(x_i), k quadratic terms x^2 and $\frac{k(k-1)}{2}$ interaction terms of the form, $x_i x_j$, where $(i < j, i, j = 1,2, \dots, k)$.

For the case of seven variables, which we are interested in this work, the second-order model is given as

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \cdots + \hat{\beta}_7 x_7 + \hat{\beta}_{11} x_1^2 + \hat{\beta}_{22} x_2^2 + \cdots + \hat{\beta}_{77} x_7^2 + \hat{\beta}_{12} x_1 x_2 + \hat{\beta}_{13} x_1 x_3 + \cdots + \hat{\beta}_{67} x_6 x_7 \tag{3}$$

which could be written in compressed form as:

Article DOI: 10.52589/AJMSS-OO1VGAGD
DOI URL: https://doi.org/10.52589/AJMSS-OO1VGAGD

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 8, Issue 3, 2025 (pp. 47-65)
www.abjournals.org

$$\hat{y} = \hat{\beta}_0 + \sum_{i=1}^{7} \beta_i x_i + \sum_{i=1}^{7} \beta_{ii} x_i^2 + \sum_{i=1}^{6} \sum_{j=2}^{7} \beta_{ij} x_i x_j \; i,j = 1,2,\ldots,7 \quad (4)$$

and $p = \frac{(k+1)(k+2)}{2} = \frac{(7+1)(7+2)}{2} = 36$ is the number of parameters for $k = 7$.

Table 1 was constructed according to the response surface model. Considering model (3) above where:

i). $x_1, x_2 \ldots, x_7$ are the main effect variable in the model;

ii). $x_1^2, x_2^2 \ldots, x_7^2$ are the second order effect variables in the model;

iii). $x_1x_2, x_1x_3, \ldots, x_6x_7$ are the interactions of the variables in the model; and $\beta_o, \beta_1, \ldots, \beta_{67}$ are the model parameters, $x_1, x_2, \ldots, x_7$ are the variables used in the construction of Table 1 which is called a design matrix, denoted by D. The extended matrix which is called the X - matrix, is constructed with the use of all the variables in equation (3). It is the X-matrix that we actually use in our analysis and in the case of $k = 7$, it is made up of 36 parameters (that is $(k+1)(k+2)/2 = 36 \; where \; k = 7$), this design has 145 runs by 36 parameters (145×36). The miniature of the X-matrix is shown on Table 2 for clarity.

Minimax Loss Criterion

Akhtar and Prescott (1986) developed a minimax loss criterion which works on weighing the effect of missing values and finding designs robust to missing observations. The X matrix constructed from design matrix according to response surface model, (see for instance eqn. (1) and the $n \times k$ design matrix, D, which may be defined as $\{x_{ui}\}$, with the elements $(x_{u1}, x_{u2}, \ldots., x_{uk})$ of the u^{th} row vector, x_{ui}, being the levels at which the predictor variables, $x_1, x_2, \ldots., x_k$, are set in the u^{th} runs of the experiment are used in the analysis of missing values. For simple illustration, let us use as an example, a design for which $k = 2, n_f = 4, n_a = 4, \; n_c = 3 \; and \; n = 11$, the matrices, D and X may be represented as follows:

Table 2: The $D \; and \; X - Matrix$ Constructed According to Response Surface Methodology

$$D = \begin{bmatrix} x_1 & x_2 \\ -1 & -1 \\ 1 & -1 \\ -1 & 1 \\ 1 & 1 \\ \alpha & 0 \\ -\alpha & 0 \\ 0 & \alpha \\ 0 & -\alpha \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \quad X = \begin{bmatrix} x_0 & x_1 & x_2 & x_1^2 & x_2^2 & x_1x_2 \\ 1 & -1 & -1 & 1 & 1 & 1 \\ 1 & 1 & -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & \alpha & 0 & \alpha^2 & 0 & 0 \\ 1 & -\alpha & 0 & \alpha^2 & 0 & 0 \\ 1 & 0 & \alpha & 0 & \alpha^2 & 0 \\ 1 & 0 & -\alpha & 0 & \alpha^2 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Article DOI: 10.52589/AJMSS-OO1VGAGD
DOI URL: https://doi.org/10.52589/AJMSS-OO1VGAGD

If there are two missing observations then two rows in the design matrix, D, and in X matrix will be missing and the remaining X is called the reduced matrix and is represented by X_r. The matrix, $X'X$, is the complete information matrix and the determinant of the complete information matrix is represented by d, that is: $d = |X'X|$

In general, if m observations are missing, then m rows in the X matrix will be missing and the matrix of m rows corresponding to the m missing observations is represented by $X_m(where\ m = 1,2, \dots)$. The $(X_r{}'X_r)$ information matrix for the reduced design will be a different matrix from the complete information matrix $(X'X)$ and its determinant will be less than the determinant of the matrix $(X'X)$ for the complete design X may be partitioned as $X = [X_m\ X_r\ ]$.

Akhtar and Prescott (1986) had developed a relation between the determinant for reduced matrix and the determinant for complete information matrix as follows: $|\,X'X| = |X_m'X_m| + |X_r'X_r||X_r'X_r| = |\,X'X| - |\,X_m'X_m|$ Ahmad, Akhtar and Gilmour (2012) discussed a relation between $\left|X_r^{'}X_r\right|$ and $|X'X|$ by using identities on the expansion of a bordered determinant as: $\left|X_r^{'}X_r\right| = |X'X|\ A_{uv\dots}$

where $A_{uv\dots}$ is the corresponding diagonal element of the m^{th} compound of the $(I - R)$ matrix and $R = X(X'X)^{-1}X'$ of order n. This means that A_u is the u^{th} diagonal element of the first compound and A_{uv} the uv^{th} diagonal element of the second compound of the $(I - R)$ matrix.

Akhtar and Prescott (1986) originated the relative reduction in $|X'X|$ due to missing observations and called it minimax loss. Let the relative loss due to any set of m missing observations out of n design points be denoted by $l_{uv\dots,}$ and be defined as in the equation below:

Relative loss : $l_{uv..} = \frac{|X'X| - |X_r'X_r|}{|X'X|} = 1 - \frac{|X_r'X_r|}{|X'X|} = 1 - A_{uv\dots} = |1 - (X_m'X_m)(X'X)^{-1}|$

The value of $l_{uv..}$ lie between zero and one i.e. $0 \le l_{uv..} \le 1$ and it has been observed that all specific types of losses are invariant to which one design point is missing and $l_{uv..}$ is constant for a specific design with fixed n and p, where n is the number of design point and p the parameters in the design, see for example Ahmad *et al.* (2012), Akram, Akhtar and Tahir (2003), Akram (2002), Akram, Akhtar and Tahir (2004) and Akhtar (2001).

If the relative loss $l_{uv..} = 1,$ then $1 - \frac{d_i}{d} = 1$ or $d - d_i = d or\ d_i = 0$

where d_i is the determinant of the reduced information matrix. This shows that the determinant of the reduced information matrix is zero, that is, $\left|X_r^{'}X_r\right|$ is singular.

If $l_{uv..} = 0,$ then $1 - \frac{d_i}{d} = 0\ \ d - d_i = 0\ \ d = d_i$ this could mean that the missing observations have no effect on the design.

Our interest is to minimize the relative loss, $l_{uv..}$ of m missing observations. The loss becomes minimum if

$A_{uv\dots} = |\,1 - (X_m'X_m)(X'X)^{-1}| = \frac{|X_r'X_r|}{|X'X|}$ tends to one.

Article DOI: 10.52589/AJMSS-OO1VGAGD
DOI URL: https://doi.org/10.52589/AJMSS-OO1VGAGD

Measure of Optimality based on the Eigenvalues of Information Matrix in Central Composite Designs (CCD)

To compare a class of ccd on the basis of optimality, a simple characterization of the optimality may be given in terms of the eigenvalues of $(X'X)$. Akram (2002) and Akram, Akhtar and Tahir (2003) stated that by using the coding convention (± 1) on the second-order polynomial equation, the eigenvalues of information matrix $(X'X)$ in ccd are:

$$\lambda_1 = \frac{\left[e + (k-1)f + n - [(e + (k-1)f + n)^2 - 4(e \times n + (k-1)f \times n - kT^2]^{\frac{1}{2}}\right]}{2}$$

$$\lambda_2 = \frac{\left[e + (k-1)f + n + [(e + (k-1)f + n)^2 - 4(e \times n + (k-1)f \times n - kT^2]^{\frac{1}{2}}\right]}{2}$$

There are k eigenvalues with $\lambda_i = T, i = 3, \ldots k+2$

$(k-1)$ eigenvalues with $\lambda_j = e - f,\ j = (k+3), \ldots, (2k+1)$

$k(k-1)/2$ eigenvalues with $\lambda_u = f,\ u = (2k+2), \ldots, p$

where $p = 1 + 2k + k(k-1)/2,\ T = n_f + 2\alpha^2$

$$e = n_f + 2\alpha^4 \ and \quad f = n_f$$

A particular member of the CCD class can be identified by using the eigenvalues of the information matrix of the model under study. Different values of α can be used to define eigenvalues for various members of the class such as cuboidal, orthogonal, rotatable, outlier robust, spherical, minimax loss1 and minimax loss2 designs: Akram, Akhtar and Tahir (2003)

Effect of Missing Observations on Trace of Information Matrix

The sum of the diagonal elements of a square matrix is called the trace of the matrix, written say $t_r(A)$: i.e. for
$A = \{a_{ij}\}$for $i, j = 1, \ldots, n;\ \ t_r(A) = a_{11} + a_{22} + \cdots + a_{nn} = \sum_{i=1}^{n} a_{ii}$
When A is not square, the trace is not defined; i.e. it does not exist.
The trace of a transposed matrix is the same as the trace of the matrix itself: $t_r(A') = t_r(A)$ ($Searle$ (1982, $chapter$ 2)

Due to a combination of m missing observations, the original well-patterned design model matrix "X" is ruined and the design loses some of its desirable properties.

Our investigation on response surface designs is confined to CCD, as they are among the most useful and versatile designs for fitting second order models. Trace criterion is considered to check the variance of estimates of all the parameters and the variance of an estimated response over a specified region.

The trace criterion for efficiency of the design is $t_r(X'X)^{-1}$, which is the sum of the variances of the estimates of all components of $\underline{\beta}$ without the factor σ^2. X_r is reduced design model matrix, due to m missing observations, that is a part of the complete design matrix X

Article DOI: 10.52589/AJMSS-OO1VGAGD
DOI URL: https://doi.org/10.52589/AJMSS-OO1VGAGD

constructed according to response surface methodology. The trace criterion for efficiency of the design with a set of missing observations is then the$t_r(X'X)^{-1}$, which is the sum of the variances of the estimates of all components of $\underline{\beta}$ except the factor σ^2: Akram (2002).

Akram (2002) further stated that, the matrix, X_m, of m missing design points may be defined as the most influential or critical combination of points for which the loss of efficiency is larger (or maximum) if they are missed. The increase in trace criterion due to combination of m missing observations may be defined as $D_{m,0} = t_r(X_r'X_r)^{-1} - t_r(X'X)^{-1}$ where $D_{m,0}$ indicate the increase in trace with m missing observations in the original matrix X without repeating any observation. In this work we shall consider increases in the trace of Central Composite Rotatable Design (CCRD) only.

RESULTS

Here, results of analysis based on various criteria discussed earlier will be considered with the view of obtaining minimax loss1 (that is, a design robust to single missing value) and minimax loss2 (which means, a design robust to missing a pair of observations). Increases in trace of (CCRD), the mean losses for missing a pair of observations, standard deviations, minimum losses and maximum losses due to missing a pair of observations shall be obtained. We shall also categorize the losses according to groups of alpha (α) that produces similar sets of losses. The obtained robust design shall be compared with other competing design in the literature stated earlier. A matlab algorithm has been developed to run the $\binom{145}{2} = 10440$ reduced matrices and carry out the minimax loss computations, in addition to other computations.
The losses due to missing the various points in the design shall be represented as shown below:

- missing a factorial point (l_f)
- missing an axial point (l_a)
- missing a centre point (l_c)
- missing a pair of factorial points (l_{ff})
- missing a pair of axial points (l_{aa})
- missing a pair of centre points (l_{cc})

Losses Due to One Missing Observation

We had earlier stated that a CCD consists of three parts, namely: factorial$(n_f or 2^k)$, axial (n_a), and centre (n_c) parts, where k are the design factors. In a seven-factor ccd with single replication of factorial and axial parts, $k = 7$, which mean that the design consist of $n_f = 2^7 or\ 128$ factorial points, $n_a = 2 * 7\ or\ 14$ axial points and $n_c = 3$ centre points, which makes it $n = 145$ design points. In the case of a single missing value, it was observed that l_a remains higher than both l_f and l_c for the whole range of $\langle$ because the axial point is fewer in number (Akram, 2002). The maximum loss due to a single missing centre observation is at $\frac{1}{n_c}(or \frac{1}{3} = 0.3333)$.The result as obtained in Table 3 indicates that there is no equality, $l_f = l_a$ for any value of α. Hence, going by Akhtar and Prescott (1986), it therefore means that for seven-factor ccd with single replication of both factorial and axial parts and three centre points,

Article DOI: 10.52589/AJMSS-OO1VGAGD
DOI URL: https://doi.org/10.52589/AJMSS-OO1VGAGD

the minimax loss design robust to a single missing value does not exist. The result of our analysis is shown on Table 3.

Table 3: Maximum Losses due to a Single Missing Observation of Factorial, Axial and Centre Points at

Different Values of ⟨ with $k = 7$

Alpha values ()	Losses due to missing a centre point (l_c)	Losses due to missing a fact. point (l_f)	Losses due to missing an axial point (l_a)
1.0	0.0754	0.2257	0.4916
1.5	0.1067	0.2247	0.4944
2.0	0.1824	0.2232	0.4917
2.0292	0.1891	0.2231	0.4912
2.5	0.3177	0.2211	0.4819
2.5729	0.3292	0.2207	0.4826
2.6	0.3317	0.2206	0.4833
$\sqrt{7} = 2.6458$	0.3333	0.2204	0.4849
3.0	0.2520	0.2189	0.5159
3.3636	0.1401	0.2176	0.5518
3.5	0.1124	0.2171	0.5623
4.0	0.0564	0.2151	0.5923

It is observed from Table 3 that the maximum loss due to missing a centre point is at $\sqrt{7} = 2.6458$. We also noticed steady increase and decrease for the range of α $values$, 1.0 to 4.0.

Losses Due To a Pair of Missing Observations

There are different combinations of a pair of missing observations. The possible combinations of two missing design points for n design points are $\binom{n}{2}$. The losses of these $\binom{n}{2}$ combinations of two missing observations fall into different groups with the same losses. The size of loss depends upon particular combinations of a pair of missing observations at different values of , where is the distance of the axial points from the design centre. There are six possible categories of combinations of three kinds of design point ($f, aandc$). In each category, all the combinations of two missing observations have the same losses in different groups.

These categories of combinations are:

(i) cc: Two center points

(ii) ac: One axial point and one centre point

(iii) fc: One factorial point and one centre point

(iv) fa: One factorial point and one axial point

Article DOI: 10.52589/AJMSS-OO1VGAGD
DOI URL: https://doi.org/10.52589/AJMSS-OO1VGAGD

(v) aa: Two axial points

(vi) ff: Two factorial points.

The table of maximum losses due to missing a pair of "cc", "ff", "aa" and "fa" design points are constructed at different values of 〈 with $k = 7$ factors, $n_f = 128$ factorial, $n_a = 14$ axial and $n_c = 3$ centre points, $n = n_f + n_a + n_c = 145$ total design points which is $\binom{145}{2} = 10440$ reduced matrices.

When any one of all the possible $(n\ 2\ )$ combinations of two observations (m_{uv}) is missing it will be one of the above mentioned combinations, where (m_{uv}) is the combination of $u^{th}$and v^{th} missing observations,

$$u = 1, 2, \dots, n-1 \quad ; \qquad v = 2, 3, \dots, n\,; \qquad\qquad u < v$$

and n is the total number of design points in CCD.

From the result of our analysis shown on Table 4, it is observed that $(l_{cc}) = \frac{m_{cc}}{n_c} = 0.6667\ at = \sqrt{7}$, where m_{cc} is the two missing centre points and n_c, the total number of centre points in the design, which is in line with Akhtar and Prescott (1986) which states that the maximum loss due to missing centre points always occur at $\alpha = \sqrt{k}$, and the value is generally put as $(l_{c\dots}) = \frac{m_{c\dots}}{n_c}\, at = \sqrt{k}$ where $m_{c\dots}$ is the combinations of m missing centre points and n_c is the number of centre points in the design.

The maximum loss reaches 1 and the design breakdown (i.e. the determinant of the information matrix equals zero) if the only set of centre points available in the design is lost. So we take at least $n_c = 3$ centre points in this design so that:

(i) The design does not get breakdown

(ii) We can compare other designs with the same configuration when there are two missing observations.

A Pair of Missing Values in Seven-Factor Central Composite Design (ccd) with Single Replication of Factorial and Axial Parts

Seven-factor CCD with a single replication of factorial and axial points consist of $n_f = 128$ factorial point, $n_a = 14$ axial points and $n_c = 3$ centre point i.e. there are a total of $n = 145$ design points. The maximum losses due to a pair of missing observations at factorial, axial and centre points are determined at different values of 〈 with $n_c = 3$.

The possible combinations of two missing design points for $k = 7$ and $n = 145$ design points are $\binom{145}{2} = 10440$ reduced matrices. The result of our analysis is as shown below on Table 4

Article DOI: 10.52589/AJMSS-OO1VGAGD
DOI URL: https://doi.org/10.52589/AJMSS-OO1VGAGD

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 8, Issue 3, 2025 (pp. 47-65)

Table 4: Maximum Losses Due to a Pair of Missing Observations at Different Values of 〈 Ranging between 1.0 and 4.0, $k = 7$,$n_f = 128$, $n_a = 14$, $n_c = 3$ **and** $p = 36$

Alpha values (〈)	Losses due to missing a pair of centre point (l_{cc})	Maximum losses due to missing:		
		a pair of factorial points (l_{ff})	a pair of axial points (l_{aa})	a factorial and an axial point (l_{fa})
1.0	0.1508	0.4144	0.9683	0.6064
1.5	0.2133	0.4130	0.9563	0.6081
2.0	0.3648	0.4110	0.9290	0.6054
2.0292	0.3783	0.4109	0.9266	0.6050
2.4500	0.6129	0.4084	0.8889	0.5973
2.5	0.6354	0.4081	0.8859	0.5970
2.5638	0.6563	0.4076	0.8837	0.5973
2.5729	0.6584	0.4076	0.8835	0.5974
2.6	**0.6634**	**0.4074**	**0.8833****	**0.5979**
2.6115	0.6648	0.4073	0.8834	0.5982
2.6458	0.6667	0.4071	0.8839	0.5991
3.0	0.5040	0.4050	0.9198	0.6226
3.3636	0.2803	0.4034	0.9603	0.6500
3.5	0.2247	0.4028	0.9697	0.6580
4.0	0.1128	0.4003	0.9877	0.6807

** Minimax loss 2

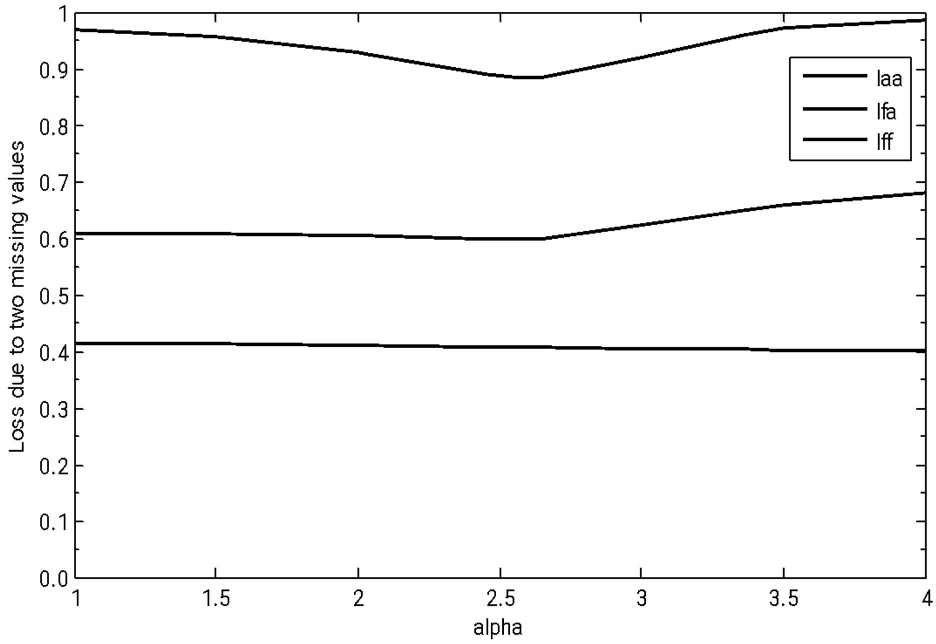
The results we obtained for maximum losses due to a pair of missing observations at "cc", "fa", "aa" and "ff" in the design for different values of alpha (〈) with $= 7n_f = 128$,$n_a = 14$, $n_c = 3$ and $n = 145$ design points are presented in Table 4, where the losses, l_{cc}, l_{ff}, $l_{aa} and l_{fa}$ are shown. Also for the design understudy and 〈 ranging from 1.0 to 4.0, the maximum loss (l_{aa}) = 0.8833 is minimum at 〈 = 2.6, so the seven-factor ccd with = 2.6 is a **Design Robust** to a pair of missing observations or say **minimax loss 2 design**.

It can be observed from Table 4 that l_{ff} is a decreasing functioning of increasing α values, while l_{aa} is an increasing and decreasing function of α

Table 4 also suggested that, for $\geq$ 1.0 the loss for missing a pair of axial points is larger than the losses for missing any other pair of factorial points.

This is because of the small number of axial points as compared to the factorial points: see Akram (2002),for more detail

 Article DOI: 10.52589/AJMSS-OO1VGAGD
DOI URL: https://doi.org/10.52589/AJMSS-OO1VGAGD



.

Figure 1: Graphical Representation of Maximum losses due to a Pair of Missing Values at various Values of ⟨

Comparison of the Minimax Loss 2 Design at ⟨ = 2.6 and $n_c = 3$ Centre Points with Designs of the Same Configuration

The following designs which we are going to use and compare with the obtained minimax loss 2 in Table 5 and its values given as: (i) Cuboidal ccd with ⟨ = 1.0

(ii) Orthogonal ccd with ⟨ = 2.0292

(iii) Rotatable with ⟨ = 3.3636

(iv) Spherical ccd with ⟨ = $\sqrt{7}$

(v) Design with minimum variance of losses or outlier robust design with ⟨ = 2.45

The maximum losses corresponding to missing a combination of two-factorial, two-axial, one-factorial and one-axial point for different designs of the same configurations are shown in Table 4

Article DOI: 10.52589/AJMSS-OO1VGAGD
DOI URL: https://doi.org/10.52589/AJMSS-OO1VGAGD

Table 5: Comparison of Maximum Losses Due to Different Pairs of Missing Observation for $k = 7$; Total Design Point $n = 145$, $n_c = 3$

Design	Apha	$\lvert X^1X \rvert$ for complete Design	$\lvert X_r' X_r \rvert$ for Reduced Design	Maxloss due to two factorial observation missing ("ff")	Maxloss due to two axial observations missing (aa)	Maxloss due to an axial and one factorial observation missing (fa)	Overall maxloss due to a pair of observations missing
Alpha = 1.0	1.000	8.531E+64	7.2463E+64	0.4144	0.9683	0.6064	0.9683
Orthogonal	2.0292	1.1613E+72	7.2200E+71	0.4109	0.9266	0.6050	0.9266
Outlier Robust	2.4500	8.3199E+73	3.2205E+73	0.4084	0.8889	0.5973	0.8889
Minimaxloss2	2.6000	3.5192E+74	1.1847E+74	0.4074	0.8833	0.5979	0.8833**
Alpha =√k	2.6458	5.4864E+74	1.8288E+74	0.4071	0.8839	0.5991	0.8839
Rotatable	3.3636	7.2626E+77	5.2269E+77	0.4034	0.9603	0.6500	0.9603

** Minimaxloss due to two missing observations

Table 5 shows that minimax loss 2 design compared with the other $ccd$of the same configuration is found to have minimum loss than any of the other such designs.

Table 6 below indicates our results as obtained for the computations of the mean, standard deviations minimum and maximum values due to missing a pair of observations.

Table 6: The Mean, Standard Deviation, Minimum and Maximum Values of the Losses Due to the

$\binom{145}{2} = 10440$ **Reduced Matrices in ccd with $k = 7$, $n_f = 128$, $n_a = 14$, $n_c = 3$ and** $n = 145$

Design Points

Alpha values (〈)	**Mean**	**Standard Deviation**	**Minimum Value**	**Maximum Value**
1.0	0.4362	0.0881	0.1508	0.9683
1.5	0.4362	0.0878	0.2133	0.9563
2.0	0.4362	0.0852	0.3648	0.9290
2.0292	0.4362	0.0850	0.3701	0.9266
2.45	0.4362	**0.0826***	0.3937	0.8889
2.5	0.4362	0.0829	0.3934	0.8859
2.5638	0.4362	0.0834	0.3929	0.8837

Article DOI: 10.52589/AJMSS-OO1VGAGD
DOI URL: https://doi.org/10.52589/AJMSS-OO1VGAGD

2.5729	0.4362	0.0834	0.3928	0.8835
2.6	0.4362	0.0838	0.3926	**0.8833****
2.6115	0.4362	0.0840	0.3924	0.8835
2.6458	0.4362	0.0844	0.3923	0.8839
3.0	0.4362	0.0936	0.3899	0.9198
3.3636	0.4362	0.1067	0.2803	0.9603
3.5	0.4362	0.1109	0.2247	0.9697
4.0	0.4362	0.1228	0.1128	0.9877

*Minimum standard deviation, **Minimax loss due to missing two observations

From Table 6 the minimum standard deviation, 0.0826, of all the losses (l_{uv}) occurred at 〈 = 2.45; thus, the design for 〈 = 2.45 is minimum variance design robust to two missing observations, for singly replicated ccd with $k = 7$, $n_f = 128$, $n_a = 14$, $n_c = 3$ and $n = 145$ design points.

Table 7: Increase in Trace Criterion due to a Combination of Two Missing Observations in $(ccrd)$ for $k = 7$ with

〈 = 3.3636, $n_f = 128$,$n_a = 14$, $n_c = 3$ and $n = 145$

S/No	**Increase in trace ($D_{2,0}$)**	**Combination of two missing Observations**	**No of Combinations**
1	0.0611	143,144 cc	3
2	0.0279	1,143 fc	384
3	0.0044	1,94 ff	5824
4	0.0045	1,4 ff	1792
5	0.0049	1,65 ff	448
6	0.0050	1,128 ff	64
7	0.0464	133,143 ac	42
8	0.0189	1,137 fa	896
9	0.0187	1,130 fa	896
10	0.0320	129,133 aa	84
11	0.2971	129,130 aa	7

Different Groups of Pairs with Similar Losses

All possible pairs of observations, $\binom{n}{2}$, can be placed in groups with similar losses. The main groups are: ff,fa ,aa ,fc ,ac where f,a and c are for a factorial, an axial and a centre point respectively.

There are three categories of 〈 producing different sets of similar losses.

For example, **Category A**: 〈 = 1.0, 1.5, 2.0, 2.0292, 2.5 and 4, they have 12 sets of similar losses as shown on Table 9 and column 2 of Table 8

Category B: 〈 = 2.5638 and 2.6458, have 13 sets of similar losses as seen in Table 10 and column three of Table 8

Article DOI: 10.52589/AJMSS-OO1VGAGD
DOI URL: https://doi.org/10.52589/AJMSS-OO1VGAGD

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 8, Issue 3, 2025 (pp. 47-65)
www.abjournals.org

Category C:〈 = 2.45, 2.5729, 2.6, 2.6115, 3.0, 3.3636, and 3.5 have 14 sets of similar losses as seen on Table 11 and column four of Table 8.

Summary of the formed groups for different values of 〈 for $k = 7$ complete replication of factorial and axial points with centre points $n_c = 3$ is shown in Table 8.

Table 8: Categories A, B, C and Different Set of Similar Losses Combination Due to Two Missing Observations

Combination of Design point	No. of Groups with equal losses		
	Category A	**Category B**	**Category C**
cc	1	1	1
ac	1	1	1
fc	1	1	1
fa	1	2	2
aa	2	2	2
ff	k-1	k-1	k
Total	**12**	**13**	**14**

Table 9: Category A with Different Sets of Similar losses, Number of Combinations and Combinations of Two Missing Observations

Alpha S/N	1.0	1.5	2.0	2.0292	2.5	4.0	No. of Combinations	Combinations of two missing observations	
1.	0.1508	0.2133	0.3648	0.3783	0.6354	0.1128	3	143,144	cc
2.	0.2841	0.3074	0.3649	0.3701	0.4685	0.2596	384	1,143	fc
3.	0.4005	0.3989	0.3966	0.3965	0.3934	0.3840	3584	1,32	ff
4.	0.4010	0.3994	0.3971	0.3970	0.3939	03845	2240	1,16	ff
5.	0.4020	0.4003	0.3980	0.3978	0.3946	0.3852	1344	1,4	ff
6.	0.4021	0.4005	0.3983	0.3981	0.3951	0.3862	448	1,96	ff
7.	0.4141	0.4123	0.4097	0.4096	0.4061	0.3959	448	1,2	ff
8.	0.4144	0.4130	0.4110	0.4109	0.4081	0.4003	64	1,128	ff
9.	0.5341	0.5534	0.5901	0.5930	0.6474	0.6174	42	131,143	ac
10.	0.6040	0.6081	0.6053	0.6050	0.5970	0.6806	1792	1,129	fa
11	0.7418	0.7449	0.7430	0.7426	0.7355	0.8338	84	131,135	aa
12	0.9683	0.9563	0.9290	0.9266	0.8859	0.8977	7	129,130	aa

It can be observed from Table 9 that:

(i) category "A" have six (6) different sets of similar losses due to missing a pair of factorial points that is, $ff_1, ff_2, ff_3, \cdots ff_6$;

ii) there are two (2) different sets of similar losses due to missing a pair of axial points; $aa_1 \& aa_2$;

Article DOI: 10.52589/AJMSS-OO1VGAGD
DOI URL: https://doi.org/10.52589/AJMSS-OO1VGAGD

iii) there is one (1) set of similar losses due to missing a factorial and an axial point (fa); and

iv) also, there is one (1) set of similar losses due to missing a factorial point and a centre point (ac).

Table 10: Category B, with Different Sets of Similar Losses, Number of Combinations and Continuation of Two Missing Observations

Alpha S/N	2.5638	2.6458	No. of combination	Combination of two missing observation	
1	0.3929	0.3923	3584	1,123	ff
2	0.3934	0.3928	2240	1,46	ff
3	0.3941	0.3935	1344	1,19	ff
4	0.3946	0.3940	448	1,64	ff
5	0.4056	0.4049	448	1,2	ff
6	0.4076	0.4071	64	1,128	ff
7	0.4765	0.4803	384	1,143	fc
8	0.5963	0.5986	896	1,129	fa
9	0.5973	0.5991	896	1,140	fa
10	0.6526	0.6566	42	129,143	ac
11	0.6563	0.6667	3	143,144	cc
12	0.7362	0.7388	84	133,137	aa
13	0.8387	0.8839	7	129,130	aa

It is observed in Table 10 that:

(i) category B has six (6) different sets of similar losses due to missing a pair of factorial points i.e. $ff_1, ff_2, \cdots ff_6$

(ii) there are two (2) different sets of similar losses due to missing a pair of axial points; $aa_1$and aa_2;

(iii) there are two (2) sets of similar losses due to missing a factorial and an axial point ($fa_1 \& fa_2$);

(iv) there is one (1) set of similar losses due to missing a factorial and a entre point (fc).

Article DOI: 10.52589/AJMSS-OO1VGAGD
DOI URL: https://doi.org/10.52589/AJMSS-OO1VGAGD

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 8, Issue 3, 2025 (pp. 47-65)

Table 11: Category C with Different Sets of Similar Losses, Number of Combination and Combination of Two Missing Observations

Alpha S/N	2.45	2.5729	2.6	2.6115	3.0	3.3636	3.5	No. of Combinations	Combinations of two missing observations	
1	0.3937	0.3928	0.3926	0.3925	0.3899	0.3879	0.3871	1344	1,111	ff
2	0.3938	0.3929	0.3927	0.3926	0.3900	0.3988	0.3872	2240	1,8	ff
3	0.3942	0.3933	0.3931	0.3930	0.3905	0.3884	0.3876	2240	1,117	ff
4	0.3950	0.3941	0.3939	0.3938	0.3912	0.3891	0.3883	1344	1,10	ff
5	0.3954	0.3945	0.3943	0.3943	0.3918	0.3898	0.3891	448	1,112	ff
6	0.4065	0.4055	0.4053	0.4052	0.4023	0.4001	0.3992	448	1,9	ff
7	0.4084	0.4076	0.4074	0.4073	0.4050	0.4034	0.4028	64	1,128	ff
8	0.4600	0.4773	0.4791	0.4797	0.4158	0.3274	0.3052	384	1,143	fc
9	0.4969	0.5970	0.5974	0.5776	0.6221	0.6496	0.6577	986	1,129	fa
10	0.5973	0.5974	0.5979	0.5982	0.6226	0.6500	0.6580	896	1,130	fa
11	0.6424	0.6532	0.6548	0.6553	0.6416	0.6191	0.6155	42	129,143	ac
12	0.6129	0.6584	0.6634	0.6648	0.5040	0.2803	0.2247	3	143,144	cc
13	0.7355	0.7364	0.7372	0.7375	0.7677	0.7996	0.8087	84	129,137	aa
14	0.8889	0.8835	0.8833	0.8834	0.9198	0.9603	0.9697	7	129,130	aa

It can also be observed from Table 11 that:

(i) category C has seven (7) different sets of similar losses due to missing a pair of factorial points that is $(ff_1, ff_2, \dots, ff_7)$;

(ii) there are two (2) different sets of similar losses due to missing a pair of axial points; aa_1 and aa_2;

(iii) there are two sets of similar losses due to missing a factorial and an axial point; fa_1 &fa_2;

(iv) there is one (1) set of similar losses due to missing a factorial and a centre point (fc).

Article DOI: 10.52589/AJMSS-OO1VGAGD
DOI URL: https://doi.org/10.52589/AJMSS-OO1VGAGD

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 8, Issue 3, 2025 (pp. 47-65) www.abjournals.org

Formation of the Combinations of Two Missing Observations

The combination of two missing observations may be any for the $\binom{n}{2}$ combination of "n" observations in a design of the selected model. The losses of the $\binom{n}{2}$ combinations of two missing observations can be categorized into different groups with similar losses. There are three portions that are well-defined in ccd as follows:

(i) $n_f = 2^k$ factorial points for complete replication and $n_f = 2^{k-1}$factorial points for half replication;

(ii) $n_a = 2k$ axial points; and

(iii) $n_c = 3$ centre points.

Six categories of the combinations are formed from the factorial, axial and centre points when a total of two observations are missed from n_f , n_a and n_c design points.

Here we are considering the case for $k = 7$ and maximum losses for each of the six categories of ff,fa ,fc ,aa , ac and cc.

There are six or seven groups of combinations of two missing factorial points with the same loss (depending on the values of 〈) in which we are presenting the group with maximum loss.

Group six or seven consists of 64 combinations of two missing factorial points producing the same losses of which 3 are arbitrarily chosen and presented here.

Group No. 6 or 7.

$$\begin{array}{cc} \text{[Point No.} & x_1x_2\,x_3x_4x_5x_6x_7 \\ \begin{array}{c} f(1) \\ f(128) \end{array} & \begin{bmatrix} -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \end{array} \quad \begin{array}{cc} \text{Point No.} & x_1x_2\,x_3x_4x_5x_6x_7 \\ \begin{array}{c} f(2) \\ f(127) \end{array} & \begin{bmatrix} -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \end{array}$$

$$\begin{array}{cc} \text{Point No.} & x_1x_2\,x_3x_4x_5x_6x_7 \\ \begin{array}{c} f(3) \\ f(126) \end{array} & \begin{bmatrix} -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \end{array} \qquad \ldots.,$$

etc.There are two groups of the same losses of axial point irrespective of the α - values, but the group with maximum losses consist of 7 combination of two missing axial points, of which 3 are arbitrarily chosen and are presented here.

Group 2

$$\begin{array}{cc} \text{Point No.} & x_1x_2\,x_3x_4x_5x_6x_7 \\ \begin{array}{c} a(129) \\ a(130) \end{array} & \begin{bmatrix} \alpha & 0 & 0 & 0 & 0 & 0 & 0 \\ -\alpha & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array} \quad \begin{array}{cc} \text{Point No.} & x_1x_2\,x_3x_4x_5x_6x_7 \\ \begin{array}{c} a(131) \\ a(132) \end{array} & \begin{bmatrix} 0 & \alpha & 0 & 0 & 0 & 0 & 0 \\ 0 & -\alpha & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

$$\begin{array}{cc} \text{Point No.} & x_1x_2\,x_3x_4x_5x_6x_7 \\ \begin{array}{c} a(133) \\ a(134) \end{array} & \begin{bmatrix} 0 & 0 & \alpha & 0 & 0 & 0 & 0 \\ 0 & 0 & -\alpha & 0 & 0 & 0 & 0 \end{bmatrix} \end{array} \qquad \ldots., \text{etc}$$

Article DOI: 10.52589/AJMSS-OO1VGAGD
DOI URL: https://doi.org/10.52589/AJMSS-OO1VGAGD

CONCLUSION

Seven-factor central composite design was studied with respect to a pair of missing observations. The central composite designs comprise of factorial, axial and center parts. In this work, we deal with all possible combinations of the missing points. The study is based on the criterion of minimizing the maximum loss, which depends mainly on the number of factors involved in the experiment, the distance of the axial point from the design centre (〈) and position of the missing points.

The work considered the case of minimizing the maximum loss when two observations are missing in a CCD with k=7. The loss of every possible combination of two missing observations was calculated using minimax loss criterion, and groups of two missing observations producing the same losses formed. The process was repeated for a range of 〈s to locate the 〈 for which the maximum loss is minimum. Such a CCD with minimum loss is termed as minimax loss 2 design.

Fully replicated seven-factor central composite design ccd with $k = 7$, $n_{f=}128$, $n_a = 14$, $n_c = 3$ and $n = 145$ is robust (or has it minimax loss 2) at $\alpha = 2.6$ with the minimax loss value 0.8833. The minimax loss occurs when two axial points (a, a) are missing. Also design robust to minimum standard deviation was obtained at $\alpha = 2.45$ with the value 0.0826. The standard deviation robust design is otherwise known as minimum variance or outlier robust design (Akram, 2002). In the course of this study, we discovered that irrespective of the value of k, when a couple of a pair of observations is missing the design will break down.

REFERENCES

Ahmad, T., Akhtar, M. and Gilmour, S. G. (2012). Multilevel Augmented Pairs Second-Order Response Surface Designs and their Robustness to Missing Data. *Communication in Statistics-Theory and Methods*, Vol. 41: 437-452.

Akhtar, M. and Prescott, P. (1986). Response Surface designs Robust to missing observations. *Communication in Statistics-Simulation and Communication*, Vol. 15(2): 345 - 363.

Akhtar, M. (2001). Five-factor Central Composite Designs Robust to a Pair of Missing Observations. *Journal of Research (Science).* Vol. 12(2): 105 - 115.

Akram, M. (2002).Central Composite Designs Robust to Three Missing Observations. *Ph.D Thesis, The Islamia University, Bahawalpur*.

Akram, M., Akhtar, M., and Tahir, M. H. (2003). Comparison of Different Central Composite Designs. *International Journal of Agriculture and Biology*, Vol. 5(4): 571 - 575.

Akram, M., Akhtar, M., and Tahir, M. H. (2004). Structure of Loss- Sensitive Combinations of Three Missing Observations in Central Composite Designs. *Journal of Pure and Applied Sciences,* Vol. 23(1): 7 – 1.

Fujiwara, K. and Matsuura, S. (2020). Prediction Variance of a Central Composite Design with Missing Observation, Communication in Statistics-Theory and Methods,49:24,6016-6031, DOI: 10.1080/03610926.2019.1625925.

Ossai, E. O. and Oladugba, A. V. (2018). The estimation of Missing values in rectangular lattice designs. Journal of Modern Applied Statistical Methods, 17(2),ep2670, doi:10.22237/jmasm/1567468800.

Rashid, F., Akbar, A. and Arshad, H. M. (2022). Effects of Missing Observations on Predictive Capability of Augmented Box-Behnken Designs, Communication in Statistics-Theory and Methods, 51:20,7225-7242, DOI: 10.1080/03610926.2021.1872633.

Searle, S. R. (1982). *Matrix Algebra Useful in Statistics*. New York: John Wiley and Son.

Yankam, B. M. and Oladugba, A. V. (2023). Robustness of Orthogonal Uniform Composite Designs against Missing data, Communication in Statistics-Theory and Methods, 52:5, 1369-1384